[redacted]

Amer. J. Bot. 68(1): 44–48. 1981.

Exhibit A

OCCURRENCE OF LOW MOLECULAR WEIGHT AND HIGH CYSTEINE CONTAINING ALBUMIN STORAGE PROTEINS IN OILSEEDS OF DIVERSE SPECIES[1]

RICHARD J. YOULE[2] AND ANTHONY H. C. HUANG

Biology Department, University of South Carolina, Columbia, South Carolina 29208

ABSTRACT

The proteins in the oilseeds of species from 11 families, including sunflower, mustard, linseed, almond, lupin, peanut, cucumber, Brazil nut, hazelnut, yucca, castor bean, and cottonseed were studied. Sucrose gradient centrifugation showed that a substantial proportion of the total seed protein from each species migrated with a 2S sedimentation coefficient. The 2S proteins, being water-soluble and thus termed albumins, comprised 20–60% of the total seed proteins, while faster migrating globulins comprised the rest. The amino acid compositions of the 2S proteins were characterisitic of storage proteins by having a high amide content. However, the 2S proteins are different from the classical globulin storage proteins in having a high content of cysteine. It is proposed that 2S albumins are seed storage proteins with a wide distribution and with chemical properties distinct from those of the globulin storage proteins. They play an additional and unique role of providing sulfur reserve for germination.

SEED PROTEINS have been traditionally divided into classes based on their solubility in different solvents. In Angiospermae with the exception of Gramineae, the seeds contain primarily globulins and albumins (Osborne, 1924; Danielsson, 1949; Derbyshire, Wright and Boulter, 1976). Globulins are insoluble in water but soluble in concentrated salt solutions, whereas albumins are soluble in water or dilute salt solution. Globulins have long been considered as storage proteins while, until recently, albumins have been thought of as metabolic protein (Danielsson, 1956; Boulter and Derbyshire, 1971; Millerd, 1975; Ashton, 1976; Youle and Huang, 1978a, 1979).

Globulins have been extensively characterized, particularly in nutritionally important legumes and oilseeds. They are generally of high molecular weight with sedimentation coefficients ranging from 7 to 13S. These proteins contain high amounts of arginine, glutamine (and glutamate) and asparagine (and aspartate). Such an amino acid composition with high nitrogen content is important since globulins are storage proteins supplying nitrogen for germination. Globulins contain a low amount of sulfur-containing amino acids, which may be a limiting factor in their nutritional value (Derbyshire et al., 1976).

Albumins are widely distributed in seeds of diverse species. Recently, we have done intensive studies on the albumins of castor bean and cottonseed (Youle and Huang, 1976, 1978a, b, 1979). The seed albumins of these two species have 2S sedimentation values and function as storage proteins in germination. Their amino acid compositions are more similar to each other than to the globulin storage proteins of the respective seeds. Other than the high amount of arginine, glutamine (glutamate), and asparagine (aspartate), these albumins are exceptionally rich in cysteine.

To see if seed albumin proteins with unique characteristics similar to those of castor bean and cottonseed are widely distributed, we examined the proteins from seeds of a diverse taxonomic array. The results, reported in this paper, reveal a distinct class of seed storage protein with a wide distribution.

MATERIALS AND METHODS—*Plant Materials*—Seeds were of the species *Lupinus polyphyllus* Lindl. (lupin) from F. W. Schumacher Co., Sandwich, Mass., *Linum usitatissimum* L. (linseed) from Schafer Seed Co., Oakes, N.D., *Cucumis sativus* L. (cucumber) from Food-Machine-Chemical Corp., Modesto, Calif., *Yucca* spp., *Brassica* spp. (mustard), *Helianthus annuus* L. (sunflower), *Bertholletia excelsa* H.B.K. (Brazil nut), *Corylus* spp. (hazelnut), *Prunus amygdalus* Batsch. (almond) and *Arachis hypogaea* L. (peanut) from local sources, *Ricinus communis* L. var. Hale (castor bean) from Baker Castor Oil Company,

[1] Received for publication 14 December 1979; revision accepted 10 April 1980.

This work was supported by NSF grant PCM 77-17679. The authors thank D. Claybrook and M. Felder for their assistance in amino acid analyses.

[2] Present address: Section of Biophysical Chemistry, Laboratory of Neurochemistry, National Institute of Mental Health, Bethesda, Maryland 20014.

Send reprint requests to A. H. C. Huang.

TABLE 1. Amount of 2S, 7 ... of various species

Family, species (common name)
Compositae *Helianthus annuus* (sunflower)
Cruciferae *Brassica* spp. (mustard)
Linaceae *Linum usitatissimum* (linseed)
Rosaceae *Prunus amygdalus* (almond)
Leguminosae *Lupinus polyphyllus* (lupin)
Arachis hypogaea (peanut)
Cucurbitaceae *Cucumis sativus* (cucumber)
Lecythidaceae *Bertholletia excelsa* (brazil nut)
Betulaceae *Corylus* spp. (hazelnut)
Liliaceae *Yucca* spp. (yucca)
Euphorbiaceae *Ricinus communis* (castor bean)
Malvaceae *Gossypium hirsutum* (cotton)

[a] 13S for peanut and c... cotton.

Plainview, Texas, an... cv. SJ-2 from Produc... California.

Protein extraction ... from the seeds by g... phosphate buffer, pl... scribed previously (... 1979).

Sedimentation su... gation—Sedimentati... performed accordir... (1961) with modific... bach, 1974) as descri... Huang, 1978a). Brief

) HIGH CYSTEINE
EINS IN

IUANG
Carolina 29208

r, mustard, linseed,
un, and cottonseed
portion of the total
it. The 2S proteins,
total seed proteins,
positions of the 2S
itent. However, the
ving a high content
a wide distribution
oteins. They play an

ting factor in their nutri-
hire et al., 1976).
ly distributed in seeds of
cently, we have done in-
e albumins of castor bean
oule and Huang, 1976,
seed albumins of these
sedimentation values and
proteins in germination.
mpositions are more sim-
an to the globulin storage
ective seeds. Other than
arginine, glutamine (glu-
gine (aspartate), these al-
nally rich in cysteine.
amin proteins with unique
ar to those of castor bean
widely distributed, we ex-
s from seeds of a diverse
ne results, reported in this
inct class of seed storage
distribution.

METHODS—*Plant Mate-*
of the species *Lupinus po-*
in) from F. W. Schumacher
s., *Linum usitatissimum* L.
fer Seed Co., Oakes, N.D.,
.. (cucumber) from Food-
Corp., Modesto, Calif.,
ca spp. (mustard), *Helian-*
nflower), *Bertholletia ex-*
zil nut), *Corylus* spp. (ha-
ygdalus Batsch. (almond)
aea L. (peanut) from local
mmunis L. var. Hale (cas-
ker Castor Oil Company,

TABLE 1. *Amount of 2S, 7S and 11S proteins in the seeds of various species*

Family, species (common name)	% of total seed protein		
	2S	7S	11S[a]
Compositae *Helianthus annuus* (sunflower)	62	0	38
Cruciferae *Brassica* spp. (mustard)	62	0	38
Linaceae *Linum usitatissimum* (linseed)	42	0	58
Rosaceae *Prunus amygdalus* (almond)	25	15	60
Leguminosae *Lupinus polyphyllus* (lupin)	38	26	36
Arachis hypogaea (peanut)	20	6	74
Cucurbitaceae *Cucumis sativus* (cucumber)	56	17	27
Lecythidaceae *Bertholletia excelsa* (brazil nut)	30	9	61
Betulaceae *Corylus* spp. (hazelnut)	28	12	60
Liliaceae *Yucca* spp. (yucca)	27	16	57
Euphorbiaceae *Ricinus communis* (castor bean)	44	14	42
Malvaceae *Gossypium hirsutum* (cotton)	33	35	32

[a] 13S for peanut and cucumber, and 9S for yucca and cotton.

Plainview, Texas, and *Gossypium hirsutum* L. cv. SJ-2 from Producer Cotton Oil Co., Fresno, California.

Protein extraction—Protein was extracted from the seeds by grinding in 0.035 M sodium phosphate buffer, pH 7.5 in 1 M NaCl as described previously (Youle and Huang, 1978a, 1979).

Sedimentation sucrose gradient centrifugation—Sedimentation ultracentrifugation was performed according to Martin and Ames (1961) with modification (Hill and Breidenbach, 1974) as described previously (Youle and Huang, 1978a). Briefly, the protein solubilized in 1 M NaCl, 0.035 M sodium phospate buffer pH 7.4, was centrifuged in a sucrose gradient from 5 to 30% sucrose in the same buffer. Myoglobin and bovine liver catalase were used as markers for sedimentation values.

Protein hydrolysis and amino acid analysis—After dialysis against water and lyophilization, the protein samples were hydrolyzed to amino acids with 4 N methanesulfonic acid in vacuo. The procedure that preserved tryptophan was followed (Simpson, Neuberger and Liu, 1976). The protein samples were reduced with dithiothreitol for half-cysteine analysis. Amino acid analyses were performed on a Beckman 120-C instrument.

RESULTS—The proteins from the seeds of species of 11 different families (Table 1) were studied. The seed proteins soluble in 1 M salt solution were extracted and analyzed by sedimentation sucrose gradient centrifugation. The sedimentation patterns of these proteins, which included albumins and globulins, are shown in Fig. 1. Two basic types of sedimentation profiles were found. Sunflower, mustard, and linseed contained primarily 2S and 11S proteins, while almond, lupin, peanut, cucumber, Brazil nut, and hazelnut had 2S, 7S and 11–13S proteins. Some seeds, such as linseed, lupin and hazelnut, contained a minor peak of protein with a sedimentation value higher than 11S. The 7S and 11S proteins of several species reported here have been well characterized as the globulin storage proteins (Derbyshire et al., 1976). The 2S proteins have received considerably less attention, but have recently been characterized in castor bean (Youle and Huang, 1978a, b), cottonseed (Youle and Huang, 1979), *Brassica* (Lönnerdal and Jansen, 1972) and lupin (Gerritsen, 1956). Reports on soybean (Hill and Breidenbach, 1974), mung bean (Ericson and Chrispeels, 1973), broad bean (Mori and Utsumi, 1979), and pea (Basha and Beevers, 1975), while concerned primarily with the 7S and 11S proteins, have revealed high amounts of low molecular weight proteins in the seed extracts.

To quantitate the relative amounts of the various proteins in each species, the areas under the peaks of the sedimentation profiles (Fig. 1) were calculated. The results of this integration are shown in Table 1. The amount of 2S protein in the total salt-soluble proteins ranges from 20% in peanut to 62% in sunflower.

The solubility of the 2S and 9–13S proteins in water, 0.05 M NaCl solution, and 0.5M NaCl solution was studied (Table 2). In each species, the 2S protein is more water-soluble than the 9–13S proteins. The high solubility of the 2S

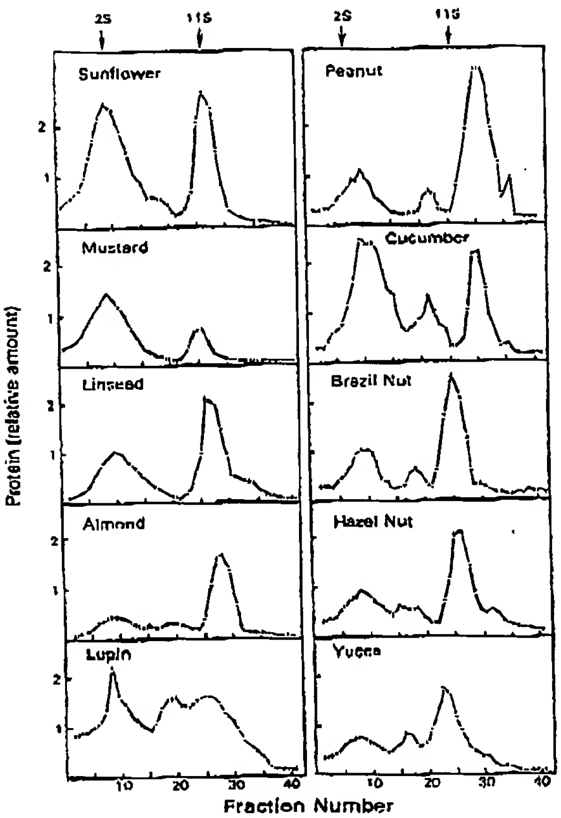

Fig. 1. Protein profiles of sucrose gradients after centrifugation of protein extracts of various seeds. The seed proteins were extracted with 1 M NaCl in 0.035 M sodium phosphate buffer, pH 7.4. Each linear gradient was composed of 35 ml of 5% (w/w) to 30% sucrose containing 1 M NaCl throughout.

TABLE 2. *Solubility of 2S and 11S proteins extracted from seeds of various species*[a]

		H_2O	0.05 M NaCl	0.5 M NaCl
Sunflower	2S	63	83	93
	11S	4	38	88
Mustard	2S	69	89	96
	11S	29	80	93
Linseed	2S	93	97	99
	11S	41	61	82
Almond	2S	88	91	92
	11S	8	67	87
Lupin	2S	57	91	91
	11S	14	40	90
Peanut	2S	88	89	93
	13S	77	82	94
Cucumber	2S	56	77	94
	13S	1	4	75
Brazil nut	2S	100	100	100
	11S	38	93	100
Hazelnut	2S	72	90	82
	11S	39	68	75
Yucca	2S	41	42	91
	9S	19	14	93
Cotton	2S	69	72	89
	9S	8	10	47

[a] The protein fractions in 1 M NaCl solution were obtained from sucrose gradients after sedimentation centrifugation. They were dialyzed against water, 0.05 M NaCl or 0.5 M NaCl, and the percent of protein remaining in the supernatant after a centrifugation of 10,000 × g for 30 min. are shown.

proteins in water is the characteristic of albumins. The 9–13S proteins are much more soluble in 0.5 M NaCl solution and thus can be classified as globulins. The 13S protein of peanut is relatively water-soluble and thus differs from the 11–13S proteins of other species.

The amino acid compositions of the 2S proteins were determined (Table 3). The 2S proteins have high nitrogen content (glutamine plus glutamate, asparagine plus aspartate, arginine). This is one of the characteristics of storage proteins, as has been established for the 7S and 9–11S globulins (Derbyshire et al., 1976). The nitrogen content of these 2S albumins is actually higher than that of the 7S and 11S globulin proteins in most species. Very significantly, the 2S proteins of the various species share a distinct similarity in their high cysteine and methionine content, which contrasts dramatically with that of the globulins (Derbyshire et al., 1976). In each species, except to a lesser degree in peanut, 6% to 13% of the total amino acids of the 2S proteins is cysteine. The methionine content ranges from 1.6 to 3.8% of the total amino acids of the 2S proteins in various species except peanut and Brazil nut. The 2S protein in peanut is very low in cysteine and methionine. Brazil nut 2S protein contains an exceptionally high amount (17%) of methionine.

DISCUSSION—Our findings clearly indicate that seed albumins of low molecular weight are abundant and occur in diverse plant species. As judged from their high nitrogen content as well as their abundance, the 2S proteins appear to be seed storage protein, serving as nitrogen reserve for germination. This idea is supported by our previous documentation of the storage role of the 2S proteins in castor bean and cottonseed (Youle and Huang, 1978a, 1979), although similar conclusive evidence for the 2S proteins in other seeds is absent. The 2S pro-

TABLE 3. *Amino acid comp*

	Sunflower	Mustard
Trp	0.28	3.60
Lys	5.31	8.70
His	1.45	3.70
Arg	5.15	4.30
Cys	6.59	9.01
Asx[a]	7.42	4.61
Thr	4.81	4.02
Ser	5.88	6.34
Glx[a]	18.36	15.89
Pro	3.94	4.50
Gly	14.18	8.48
Ala	5.06	5.26
Val	5.08	4.24
Met	2.84	2.60
Ile	4.01	3.78
Leu	5.76	6.89
Tyr	1.89	1.88
Phe	2.00	2.20

[a] Asx: aspartate and aspa
[b] +: trace amounts.

teins represent a distin
protein, with charac
those of other classes
These characteristics
weight, high solubility
teine content, in additi
nitrogen content. Alt
low molecular weight
casionally in a few i
(Gerritsen, 1956; Lönr
Tully and Beevers, 1
1978a, 1979), our syst
study establishes that
a distinct class of stor

Since our study de
proteins of the seeds, t
from some plant spe
contaminated by mino
acteristics, such as the
tease inhibitors (Lien
These contaminants
stantial amount in see
is revealed by their a
(Liener and Kakade,
ferent from those of t

The 2S protein wit
methionine content re
form in seed. During
is mobilized and utili
acid and protein struct
thesis of important co
and membrane sulfoli
no other sulfur-storage
described, and the hi
2S proteins, as comp
globulins, d cuments
tance.

"2S and 11S proteins extracted from
pecies"

	H_2O	0.05M NaCl	0.5M NaCl
	63	83	93
	4	38	88
	69	89	96
	29	80	93
	93	97	99
	41	61	82
	88	91	92
	8	67	87
	57	91	91
	14	40	90
S	88	89	93
S	77	82	94
S	56	77	94
S	1	4	75
S	100	100	100
S	38	93	100
:S	72	90	82
iS	39	68	75
:S	41	42	91
/S	19	14	93
2S	69	72	89
9S	8	10	47

:tions in 1 M NaCl solution were ob-
gradients after sedimentation centrif-
dialyzed against water, 0.05 M NaCl
l the percent of protein remaining in
r a centrifugation of 10,000 × g for 30

ally with that of the globulins
l., 1976). In each species, ex-
degree in peanut, 6% to 13%
no acids of the 2S proteins is
iethionine content ranges from
he total amino acids of the 2S
ous species except peanut and
: 2S protein in peanut is very
and methionine. Brazil nut 2S
s an xceptionally high amount
onine.

—Our findings clearly indicate
ins of low molecular weight are
occur in diverse plant species.
n their high nitrogen content as
undance, the 2S proteins appear
rage protein, serving as nitrogen
mination. This idea is supported
us documentation of the storage
proteins in castor bean and cot-
le and Huang, 1978a, 1979), al-
r conclusive evidence for the 2S
h r seeds is absent. The 2S pro-

TABLE 3. *Amino acid composition (% composition) of seed 2S proteins from various species*

	Sunflower	Mustard	Linseed	Lupin	Peanut	Cucumber	Brazil nut	Hazelnut	Yucca	Castor	Cotton
Trp	0.28	3.60	0.76	+[b]	+	0.58	+	+	0.64	+	1.68
Lys	5.31	8.70	5.96	6.61	4.03	3.30	0.89	3.42	4.12	3.0	7.46
His	1.45	3.70	1.23	2.90	2.92	1.26	1.90	0.91	2.03	0.8	2.26
Arg	5.15	4.30	6.03	6.89	7.37	11.90	11.78	11.82	11.46	9.4	10.32
Cys	6.59	9.01	8.17	6.15	1.81	8.92	13.11	10.27	9.29	8.5	7.74
Asx[a]	7.42	4.61	6.36	7.82	9.55	5.30	4.27	6.59	7.34	4.4	8.57
Thr	4.81	4.02	3.58	4.11	4.54	3.96	+	2.73	4.43	1.4	2.86
Ser	5.88	6.34	6.05	8.91	7.25	6.28	6.29	3.69	4.69	11.8	4.05
Glx[a]	18.36	15.89	23.75	24.71	16.74	20.02	24.53	29.69	19.34	30.0	27.23
Pro	3.94	4.50	1.57	3.25	3.77	3.07	4.63	1.27	4.56	2.5	2.96
Gly	14.18	8.48	13.79	7.38	8.38	10.89	5.71	7.25	8.41	8.4	8.69
Ala	5.06	5.26	5.14	4.59	5.64	4.94	1.33	3.92	5.33	3.7	4.39
Val	5.08	4.24	3.85	3.15	6.62	2.83	0.52	2.78	4.12	4.6	1.68
Met	2.84	2.60	1.93	0.66	1.84	3.79	17.33	3.69	3.17	1.6	1.80
Ile	4.01	3.78	2.87	2.43	3.78	2.22	0.90	2.74	2.19	3.0	1.53
Leu	5.76	6.89	5.32	5.01	6.78	6.44	5.58	5.27	4.14	4.2	2.55
Tyr	1.89	1.88	1.47	1.97	3.29	2.06	1.07	3.03	3.29	1.8	2.86
Phe	2.00	2.20	2.16	3.44	5.69	1.69	0.17	0.94	1.44	0.9	1.55

[a] Asx: aspartate and asparagine; Glx: glutamate and glutamine.
[b] +: trace amounts.

teins represent a distinct class of seed storage protein, with characteristics distinct from those of other classes of seed storage protein. These characteristics include low molecular weight, high solubility in water, and high cysteine content, in addition to the extremely high nitrogen content. Although seed proteins of low molecular weight have been reported occasionally in a few individual plant species (Gerritsen, 1956; Lönnerdal and Jansen, 1972; Tully and Beevers, 1976; Youle and Huang, 1978a, 1979), our systematic and comparative study establishes that these proteins represent a distinct class of storage protein.

Since our study deals only with the major proteins of the seeds, the 2S proteins extracted from some plant species unavoidably were contaminated by minor proteins of other characteristics, such as the well-characterized protease inhibitors (Liener and Kakade, 1969). These contaminants are not present in substantial amount in seeds and their low amount is revealed by their amino acid compositions (Liener and Kakade, 1969) which are very different from those of the 2S protein (Table 3).

The 2S protein with its high cysteine and methionine content represents a sulfur-storage form in seed. During germination, the sulfur is mobilized and utilized not only for amino acid and protein structure, but also for the synthesis of important cofactors and coenzymes, and membrane sulfolipid. To our knowledge, no other sulfur-storage form in seeds has been described, and the high sulfur content of the 2S proteins, as compared with that of 7–11S globulins, documents their relative importance.

The 2S proteins with their high cysteine content should be more valuable for human nutrition than other storage proteins (globulin) in the same seeds. Selection of plant varieties for amino acid composition of high nutritional value does not usually generate new protein species but merely shifts the relative proportions of proteins already present so that the more nutritionally valuable protein species constitute a high proportion of the total protein. The 2S protein should be an important consideration in the selection of plant varieties with a higher content of the essential amino acids, cysteine and methionine. Since we established earlier (Youle and Huang, 1978b, 1979) that the castor bean and cotton 2S proteins are potent allergens (Spies et al., 1951), the allergenic properties of the 2S proteins deserve consideration in the nutritional evaluation of this class of seed proteins.

LITERATURE CITED

ASHTON, F. M. 1976. Mobilization of storage proteins in seeds. Ann. Rev. Plant Physiol. 27: 95–117.

BASHA, S. M. M., AND L. BEEVERS. 1975. The development of proteolytic activity and protein degradation during germination. Planta 124: 77–87.

BOULTER, D., AND E. DERBYSHIRE. 1971. Taxonomic aspect of the structure of legume proteins. *In* J. B. Harborne, D. Boulter, and B. L. Turner [eds.], Chemotaxonomy of the Leguminosae, p. 285–308. Academic Press Inc., New York.

DANIELSSON, C. E. 1949. Seed globulins of the Gramineae and Leguminosae. Biochem. J. 44: 387–400.

———. 1956. Plant proteins. Ann. Rev. Plant Physiol. 7: 215–236.

DERBYSHIRE, E., D. J. WRIGHT, AND D. BOULTER. 1976. Legumin and vicilin, storage proteins of legume seeds. Phytochemistry. 15: 3–24.

ERICSON, M. C., AND M. J. CHRISPEELS. 1973. Isolation and characterization of glucosamine-containing storage proteins from the cotyledons of *Phaseolus aureus*. Plant Physiol. 52: 98–104.

GERRITSEN, T. 1956. Lupin seed proteins. IV. Amino acid composition of the globulins from *Lupinus angustifolius* and *Lupinus luteus*. Biochem. Biophys. Acta. 22: 269–273.

HILL, J. E., AND R. W. BREIDENBACH. 1974. Proteins of soybean seeds. I. Isolation and characterization of the major components. Plant Physiol. 53: 742–746.

LIENER, I. E., AND M. L. KAKADE. 1969. Protease inhibitors. *In* I. E. Liener [ed.], Toxic consituents of plant foodstuffs, p. 7–68. Academic Press Inc., New York.

LÖNNERDAL, B., AND J. C. JANSEN. 1972. Studies on *Brassica* seed proteins. I. The low molecular weight proteins in rapeseed. Isolation and characterization. Biochem. Biophys. Acta. 278: 175–183.

MARTIN, R. G., AND B. N. AMES. 1961. A method for determining the sedimentation behavior of enzymes. Application to protein mixtures. J. Biol. Chem. 236: 1372–1379.

MILLERD, A. 1975. Biochemistry of legume seed proteins. Ann. Rev. Plant Physiol. 26: 53–72.

MORI, T., AND S. UTSUMI. 1979. Purification and properties of storage proteins of broad bean. Agric. Biol. Chem. 43: 577–583.

OSBORNE, T. B. 1924. The vegetable proteins. 2nd ed., Longsman-Green, London.

SIMPSON, R. J., M. R. NEUBERGER, AND T. Y. LIU. 1976. Complete amino acid analysis of proteins from a single hydrolysate. J. Biol. Chem. 251: 1936–1940.

SPIES, J. R., F. J. COULSON, D. C. CHAMBERS, H. S. BERTON, H. STEVENS, AND J. H. SHIMP. 1951. The chemistry of allergens XI. Properties and composition of natural proteoses isolated from oilseeds and nuts by the CS-1A procedure. J. Amer. Chem. Soc. 73: 3995–4001.

TULLY, R. E., AND H. BEEVERS. 1976. Protein bodies of castor bean endosperm. Isolation, fractionation, and the characterization of protein components. Plant Physiol. 58: 710–716.

YOULE, R. J., AND A. H. C. HUANG. 1976. Protein bodies from the endosperm of castor bean. Subfractionation, protein components, lectins, and changes during germination. Plant Physiol. 58: 703–709.

———, AND ———. 1978a. Albumin storage proteins in the protein bodies of castor bean. Plant Physiol. 61: 13–16.

———, AND ———. 1978b. Identification of the castor bean allergens as the albumin storage proteins in the protein bodies of castor bean. Plant Physiol. 61: 1040–1042.

———, AND ———. 1979. Albumin storage proteins and allergens in cottonseed. J. Agric. Food. Chem. 27: 500–503.